

GenCore version 5.1.6
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: November 17, 2003, 20:18:19 ; Search time 2564 Seconds
(without alignments)
1626.292 Million cell updates/sec

```
Title: US-09-894-633A-81
Perfect score: 666
Sequence: 1 ctaccgctctatcacact.....aacggccgcgggcgttggg 666
```

Scoring table: OLIGO_NJC
Gapop 60.0 , Gapext 60.0

Searched: 2888711 seqs, 20454813386 residues

Word size :

Total number of hits satisfying chosen parameters: 5777422

```
Minimum DB seq length: 0
Maximum DB seq length: 200000000
```

Post-processing: Listing first 45 summaries

Database : GenEmbl: *

1: gb ba: *
2: gb htg: *
3: gb in: *
4: gb cn: *
5: gb cv: *
6: gb pat: *
7: gb htg: *
8: gb pl: *
9: gb pr: *
10: gb ro: *
11: gb sts: *
12: gb sy: *
13: gb un: *
14: gb vi: *
15: em ba: *
16: em fun: *
17: em hum: *
18: em in: *
19: em mu: *
20: em cm: *
21: em or: *
22: em ov: *
23: em pat: *
24: em ph: *
25: em pl: *
26: em ro: *
27: em sts: *
28: em un: *
29: em vi: *
30: em htg hum: *
31: em htg inv: *
32: em htg other: *
33: em htg mus: *
34: em htg plu: *
35: em htg rod: *
36: em htg man: *
37: em htg vit: *
38: em sy: *
39: em htg hum: *
40: em htg mus: *
41: em htg other: *

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query			DB	ID	Description
		Match	Length	%			
1	666	100.0	666	6	AX356287	Sequence	
2	112	16.8	7142	8	AF451895	Zea mays	
C 3	30	4.5	127652	2	AP001526	Homo sapi	
C 4	30	4.5	132948	2	AC006410	Homo sapi	
C 5	30	4.5	154841	2	AC093034	Homo sapi	
C 6	30	4.5	155736	2	AC004229	Homo sapi	
7	26	3.9	1092	6	AX653891	Sequence	
8	26	3.9	90817	10	AL928648	Mouse DNA	
9	26	3.9	120926	2	AC133008	Oryza sat	
10	26	3.9	155185	2	AC023549	Homo sapi	
11	26	3.9	157834	2	AC022931	Homo sapi	
12	26	3.9	164526	9	AC114969	Homo sapi	
C 13	26	3.9	183011	2	BX293990	Mus muscu	
C 14	26	3.9	190338	10	AC129216	Mus muscu	
C 15	26	3.9	198873	10	AL928914	Mouse DNA	
C 16	26	3.9	266660	10	AC122356	Mus muscu	
17	26	3.9	225997	2	AC105463	Rattus no	
C 18	26	3.9	221389	2	AC112122	Rattus no	
C 19	26	3.9	221951	10	AC122857	Mus muscu	
C 20	26	3.9	235511	2	AC121887	Mus muscu	
C 21	26	3.9	255090	2	AC098334	Rattus no	
C 22	26	3.9	285405	2	AC099437	Rattus no	
23	26	3.9	273204	2	AC119488	Rattus no	
C 24	26	3.9	275319	2	AC109121	Rattus no	
25	26	3.9	330828	2	AC098399	Rattus no	
26	25	3.8	480	14	SHU27486	U27486 Suid herpes	
C 27	25	3.8	110753	9	AC01216	Homo sapi	
C 28	25	3.8	118436	8	AP005064	Oryza sat	
C 29	25	3.8	142986	2	AC023467	Homo sapi	
C 30	25	3.8	163584	2	AC008470	Homo sapi	
31	25	3.8	189911	2	AC111975	Rattus no	
32	25	3.8	225379	2	AC121485	Rattus no	
33	25	3.8	248933	2	AC109773	Rattus no	
C 34	25	3.8	293105	2	AC094445	Rattus no	
C 35	25	3.8	250169	2	AC120668	Rattus no	
36	25	3.8	261502	2	AC094923	Rattus no	
37	24	3.6	552	3	AY113570	Drosophil	
38	24	3.6	1923	8	BM159446	Brassica na	
C 39	24	3.6	3200	3	AF095032	Drosophi	
C 40	24	3.6	6062	1	AF139177	Streptoc	
41	24	3.6	13028	2	AL237126_4	Continuation of	
42	24	3.6	17528	2	AC020192	Drosophil	
43	24	3.6	31741	9	AL607065	Human DNA	
C 44	24	3.6	46630	3	DMC115C2	Drosophil	
C 45	24	3.6	60135	2	AC091200	Drosophil	

ALIGNMENTS

	RESULT 1					PAT 06-FEB-2002
	AX356287	AX356287	Sequence al from Patent WO020905.	666 bp	DNA linear	
	LOCUS					
	DEFINITION					
	ACCESSION	AX356287				
	VERSION	AX356287.1	GI : 9620794			
	KEYWORDS	.				
	SOURCE	Zea mays				
	Taxonomy	Zea mays				
	ORGANISM	Zea mays				

REFERENCE	AUTHORS	TITLE
1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100

Conner, T.W., Dubois, P., Malven, M. and Masucci, J.D.
Plant regulatory sequences for selective control of gene expression

[illegible]

REFERENCE

2 (bases 1 to 127652)
 Hattori,M., Ishii,K., Toyoda,A., Taylor,T.D., Hong-Seog,P.,
 Fujiyama,A., Yada,T., Totoki,Y., Watanabe,H. and Sakaki,Y.
 Direct Submission
 TITLE
 JOURNAL
 Submitted (22-MAR-2000) Masahira Hattori, The Institute of Physical
 and Chemical Research (RIKEN), Genomic Sciences Center (GSC),
 Kitasato Univ. 1-15-1 Kitasato, Sagamihara, Kanagawa 228-8555,
 Japan [E-mail:hattori@gsc.riken.go.jp,
 URL:http://hgp.gsc.riken.go.jp/, Tel:81-42-778-9923,
 Fax:81-42-778-9924]

COMMENT

On May 30, 2000 this sequence version replaced gi:732854c.
 ----- Genome Center
 Center: RIKEN Genomic Sciences Center(GSC)
 Web site: http://hgp.gsc.riken.go.jp/
 Contact: hattori@gsc.riken.go.jp
 ----- Project Information
 Center project name: HumDraft11
 Center clone name: RP11-727C13
 ----- Summary Statistics
 Sequencing vector: PCR products; 100% of reads
 Chemistry: Dye-terminator ET-amersham; 100% of reads
 Assembly program: Phrap; version 0.990329
 Consensus quality: 115772 bases at least Q40
 Consensus quality: 121432 bases at least Q30
 Consensus quality: 123817 bases at least Q20
 Insert size: 125152; sum-of-contigs
 Quality coverage: 6.02x in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently consists of
 24 contigs. The true order of the pieces is not known and their
 order in this sequence record is arbitrary. Gaps between the
 contigs are represented as runs N, but the exact sizes of the gaps
 are unknown. This record will be updated with the finished sequence
 as soon as it is available and the accession number will be
 preserved.

```

1 12210 contig of 12210 bp in length
12311 30914 contig of 18604 bp in length
31015 46674 contig of 15660 bp in length
46775 56057 contig of 9283 bp in length
56158 66930 contig of 10773 bp in length
67031 73289 contig of 6259 bp in length
73390 80253 contig of 6864 bp in length
80354 85307 contig of 4954 bp in length
85408 88728 contig of 3321 bp in length
88829 90468 contig of 1539 bp in length
90468 92766 contig of 2299 bp in length
92867 96711 contig of 3845 bp in length
96812 101444 contig of 4633 bp in length
101545 105387 contig of 3843 bp in length
105488 108938 contig of 3351 bp in length
108939 112138 contig of 3200 bp in length
112139 114338 contig of 1999 bp in length
114338 117279 contig of 2942 bp in length
117280 119778 contig of 2298 bp in length
119778 123392 contig of 2615 bp in length
123392 123862 contig of 1370 bp in length
123862 125178 contig of 1216 bp in length
125178 126524 contig of 1246 bp in length
126524 127652 contig of 1028 bp in length
Sequence updated (26-May-2000).
```

NOTE: This is a 'working draft' sequence. It currently
 consists of 24 contigs. The true order of the pieces
 is not known and their order in this sequence record is
 arbitrary. Gaps between the contigs are represented as
 runs of N, but the exact sizes of the gaps are unknown.
 This record will be updated with the finished sequence
 as soon as it is available and the accession number will
 be preserved.

```

1 12210: contig of 12210 bp in length
12211 12310: gap of 100 bp
12311 30914: contig of 18604 bp in length
30915 31014: gap of 100 bp
```

```

* 31015 46674: contig of 15660 bp in length
* 46675 56057: gap of 100 bp
* 56058 66930: contig of 9283 bp in length
* 66931 73289: contig of 10773 bp in length
* 73290 80253: gap of 100 bp
* 80254 85307: contig of 6259 bp in length
* 85308 85408: contig of 6864 bp in length
* 85409 88728: gap of 100 bp
* 88729 90468: contig of 3321 bp in length
* 90469 92766: contig of 1539 bp in length
* 92767 92866: gap of 100 bp
* 92867 96711: contig of 2299 bp in length
* 96712 96811: gap of 100 bp
* 96812 101444: contig of 3845 bp in length
* 101445 101544: gap of 100 bp
* 101545 105387: contig of 4633 bp in length
* 105388 105488: gap of 100 bp
* 105489 108938: contig of 3843 bp in length
* 108939 108938: gap of 100 bp
* 108939 112138: contig of 3351 bp in length
* 112139 112138: gap of 100 bp
* 112139 114338: contig of 3200 bp in length
* 114339 114337: contig of 100 bp
* 114338 117279: contig of 1999 bp in length
* 117280 117379: gap of 100 bp
* 117380 119677: contig of 2942 bp in length
* 119678 119777: gap of 100 bp
* 119778 123392: contig of 2615 bp in length
* 123393 123492: gap of 100 bp
* 123493 123862: contig of 1370 bp in length
* 123863 123962: gap of 100 bp
* 123963 125178: contig of 1216 bp in length
* 125179 125278: gap of 100 bp
* 125279 126524: contig of 100 bp
* 126525 126624: gap of 100 bp
* 126625 127652: contig of 1028 bp in length.

FEATURES
      source
      1..127652
         /organism="Homo sapiens"
         /mol_type="genomic DNA"
         /db_xref="taxon:9606"
         /map="11q13"
         /clone="RP11-727C13"
      misc_feature
      1..12210
         /note="assembly_fragment"
      misc_feature
      12311..30914
         /note="assembly_fragment"
      misc_feature
      31015..46674
         /note="assembly_fragment"
      misc_feature
      46775..56057
         /note="assembly_fragment"
      misc_feature
      56158..66930
         /note="assembly_fragment"
      misc_feature
      67031..73289
         /note="assembly_fragment"
      misc_feature
      73390..80253
         /note="assembly_fragment"
      misc_feature
      80354..85307
         /note="assembly_fragment c:one_end:T7 vector_side:left"
      misc_feature
      85408..88728
         /note="assembly_fragment"
      misc_feature
      88829..90367
         /note="assembly_fragment c:one_end:S26 vector_side:right"
      misc_feature
      90468..92766
         /note="assembly_fragment"

```

```

misc_feature 92867..96711
              /note="assembly_fragment"
misc_feature 96812..101444
              /note="assembly_fragment"
misc_feature 101545..105387
              /note="assembly_fragment"
misc_feature 105488..108838
              /note="assembly_fragment"
misc_feature 108939..112138
              /note="assembly_fragment"
misc_feature 112239..114237
              /note="assembly_fragment"
misc_feature 114338..117279
              /note="assembly_fragment"
misc_feature 117380..119677
              /note="assembly_fragment"
misc_feature 119778..122392
              /note="assembly_fragment"
misc_feature 122493..123862
              /note="assembly_fragment"
misc_feature 123963..125178
              /note="assembly_fragment"
misc_feature 125279..126524
              /note="assembly_fragment"
misc_feature 126625..127652
              /note="assembly_fragment"

```

BASE COUNT 32152 a 31281 c 30169 g 31749 t 2301 others

Query Match 4.5%; Score 30; DB 2; Length 127652;

Best Local Similarity 100.0%; Pred. No. 0.00019;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 219 CTCGCTCTGCGCTCTCCCTCTCCCTCTCCCTCTCC 248
DB 73313 CTCGCTCTGCGCTCTCCCTCTCCCTCTCCCTCTCC 79284

```

RESULT 4

AC006410/c

LOCUS

DEFINITION Homo sapiens chromosome 11 clone PAC pJ59L21 map 11q2.2. ***

SEQUENCING IN PROGRESS ***, 6 unordered pieces.

AC006410

AC006410.2 GI:21389726

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Evans,G.A., Athanasiou,M., Aguayo,P., Arenate,S., Armstrong,D.,

Basit,M., Buettner,J., Butler,C., Card,P., deSailboat,C., Dunn,J.,

English,C., Ethridge,S., Garner,H.R., Gordon,M., Grant,O.,

Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T.,

McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,C.,

Schultz,R.A., Stimson,S., Wagner,N., Waller,M., Ward,T. and

Whitaker,T.

HTGS Submission

Unpublished

2 (bases 1 to 132948)

Evans,G.A., Athanasiou,M., Aguayo,P., Arenate,S., Armstrong,D.,

Basit,M., Buettner,J., Butler,C., Card,P., deSailboat,C., Dunn,J.,

English,C., Ethridge,S., Garner,H.R., Gordon,M., Grant,O.,

Hahner,L., Joslin,J., Lewis,E., Loo,H., Loo,K.N., Major,T.,

McFarland,J., Newton,J., Osborne-Lawrence,S., Schageman,C.,

Schultz,R.A., Stimson,S., Wagner,N., Waller,M., Ward,T. and

Whitaker,T.

Direct Submission

Submitted (16-JAN-1999)

Genome Science & Technology Center,

University of Texas Southwestern Medical Center, 5323 Harry Hines

Blvd., Dallas, TX 75235-8591, USA

On Jun 21, 2002 this sequence version replaced gi:4160136.

CHROMOSOMAL LOCUS: This PAC clone comes from the Chromosome 11q12.2
Best's disease region mapped between STS D11S461 and EST AHNAK.

This region spans over 1.5 Mbp.

MAPPER CONFIRMATION: WI-17055

MAPPED CLONE OVERLAP: PACs pJ57L14 and pJ57L28.

* NOTE: This is a 'working draft' sequence. It currently
* consists of 6 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

```

* 1 2186: contig of 2186 bp in length
* 2187 2286: gap of unknown length
* 2287 12644: contig of 10358 bp in length
* 12645 12744: gap of unknown length
* 12745 18986: contig of 6242 bp in length
* 18987 19086: gap of unknown length
* 19087 35290: contig of 16204 bp in length
* 35291 35390: gap of unknown length
* 35391 66291: contig of 30901 bp in length
* 66292 66391: gap of unknown length
* 66392 132948: contig of 66557 bp in length.

```

FEATURES

Source

```

1..132948
   /organism="Homo sapiens"
   /mol_type="genomic DNA"
   /db_xref="taxon:9606"
   /chromosome="11"
   /map="11q12.2"
   /clone="PAC pJ59L21"

```

BASE COUNT 34604 a 34236 c 31860 g 31745 t 503 others

ORIGIN

Query Match 4.5%; Score 30; DB 2; Length 132948;

Best Local Similarity 100.0%; Pred. No. 0.00019;

Matches 30; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

QY 219 CTCGCTCTGCGCTCTCCCTCTCCCTCTCCCTCTCC 248
DB 116441 CTCGCTCTGCGCTCTCCCTCTCCCTCTCCCTCTCC 116412

```

RESULT 5

AC090384

LOCUS

DEFINITION Homo sapiens chromosome 11 clone RP11-727C13 map 11, WORKING DRAFT

SEQUENCE, 13 unordered pieces.

AC090384

AC090384.3 GI:15290875

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

Birren,B., Linton,C., Nusbaum,C. and Lander,E.

Homo sapiens chromosome 11, clone RP11-727C13

Unpublished

2 (bases 1 to 154841)

Birren,B., Linton,C., Nusbaum,C., Lander,E., Allen,N., Anderson,S.,

Barna,N., Basten,J., Boguslavsky,L., Buckgalter,B., Brown,A.,

Carata,J., Campopiano,A., Choepel,Y., Colangelo,M., Collins,S.,

Collamore,A., Cooke,P., DeArelano,K., Dewar,K., Diaz,J.S.,

Dodge,S., Fero,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J.,

Gardyna,S., Girdle,S., Goyette,M., Graham,L., Grand-Pierre,N.,

Janes,C., Karatas,A., Lakoque,K., Lamazares,R., Landers,T.,

Lehoczky,J., Levine,R., Liu,G., Maclean,C., Macdonald,P.,

Marquis,N., Matthews,C., McCarthy,M., McSwan,P., McKernan,K.,

McPheters,R., Meldrum,J., Mereus,L., Mihova,T., Menga,V.,

Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H.,

COMMENT:

Submitted (26-FEB-1998) Genome Science & Technology Center,
University of Texas Southwestern Medical Center, 5323 Harry Hines
Blvd, Dallas, TX 75235-8591, USA

- * NOTE: this is a 'working draft' sequence. It currently
- * consists of 29 contigs. The true order of the pieces

* arbitrary. Gaps between the contigs are represented as

- * This record will be updated with the finished sequence
- * as soon as it is available and the accession number will

* 1 2239: contig of 2239 bp in length

- * 2340 4926: contig of 2587 bp in length
- * 4927 5026: gap of 100 bp

```

*      5027      8003: contig of 2977 bp in length:
*      8004      8103: gap of 100 bp

```

* 10511: gap of 100 bp
10510: gap of 100 bp
10510: gap of 2407 bp all ranges
8204

* 12842 12941: gap of 100 bp
* 12843 15005: contig of 2054 bp in length

* 16006 16105: gap of 100 bp
* 16106 18268: coding of 2163 bp in insert

* 18269 18368: gap of 100 bp
* 18369 21525: contig of 3157 bp in length

* 21626 2492: contig of 2867 bp in length

* 24593 27337: contig of 2745 bp in length

* 27438 30341: contig of 2904 bp in length

```

*
*
*
30442
34747: contig of 4306 bp in length
34748
34847: gap of 100 bp

```

* 34848 39231: contig of 4384 bp in length
* 39232 39331: gap of 100 bp

* 43438 gap of 100 bp
43537: contig of 4100 bp in region
43538: gap of 100 bp

* 47030 47129: gap of 100 bp
* 47130 53584: contig of 6455 bp in length

```

*
53585 53684: gap of 100 bp
*
53685 58032: contig of 4348 bp in length

```

* 58033 58132: gap of 100 bp
* 58133 63938: contig of 5806 bp in length

* 1

- * 70330 76231: contig of 5902 bp in length
- * 76232 76331: gap of 100 bp

* 76332 81290: contig of 4959 bp in length
* 81291 81390: gap of 100 bp

81392 88766: contig of 3/6 bp in region
88767 88866: gap of 100 bp

* 98894 98893: concy of 10027 bp in region
98893: gap of 100 bp

* 106977 107076: gap of 100 bp
* 107077 112508: contig of 5520 bp in length

- * 112609 112708: gap of 100 bp
- * 112709 121290: contig of 8582 bp in length

- * 121291 122395: gap of 100 bp
- * 121391 134773: contig of 13380 bp in Jera2-b

* 134771 134870: gap of 100 bp
* 134871 147313: contig of 12443 bp in length

* 147414 155736: contig of 8323 bp in length.

5.

2003, 2005) and the Mouse Genome Sequencing Consortium whole genome

shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: <http://www.sanger.ac.uk>
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases:

Em: EMBL; Sw: SWISSPROT; Tr: TREMBL; Wp: WORMPEP; Information

on the WORMPEP database can be found at

http://www.sanger.ac.uk/projects/C_elegans/wormpep RP23-379A6 is

from the RP23-3 mouse BAC library

constructed by the group of Pieter de Jong.

For further details see <http://www.chori.org/bacpac/home.htm>

VECTOR: pBAC3.6.

FEATURES

source

BASE COUNT 26550 a 19257 c 18758 g 26272 t
ORIGIN

Query Match 3.9%; Score 26; DB 10; Length 90837;
Best Local Similarity 100.0%; Pred.No. 0.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 CTCGCGCCCTCTCCCTCTCCCTCTCC 248

Db 7515 CTCGCGCCCTCTCCCTCTCCCTCTCC 7540

RESULT 9
AC133008 120926 bp DNA linear HTG 28-SEP-2002
LOCUS OSJNB0026K20, ** SEQUENCING IN PROGRESS **, 8 ordered pieces.
DEFINITION AC133008

AC133008.2 GI:23334734

HTG; HTGS_PHASE2.

ORYZA SATIVA (japonica cultivar-group)

SOURCE ORYZA SATIVA (japonica cultivar-group)

ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;

ERHARTOIDEAE; Oryzaceae; Oryza.

1 (bases 1 to 120926)

Buell, C.R., Yuan, Q., Ouyang, S., Liu, J., Gansterger, K., Jones, K.M.,

Overton II, J., Tsitiriz, T., Kim, M., Bera, J., Jit, S., Fadtosh, D.W.,

Tailon, L., Koo, H., Zismann, V., Hsiao, J., Blunt, S., Vanaken, S.,

Riedmuller, S.B., Utterbach, T., Feldblyum, T., Yang, Q., Haas, B.,

Suh, B., Peterson, C., Quackenbush, J., White, O., Salzberg, S. and

Fraser, C.

ORYZA SATIVA ssp. japonica cv. Nipponbare OSJNB0026K20 BAC genomic

TITLE

JOURNAL REFERENCE

Unpublished
2 (bases 1 to 120926)

Buell, R.

Direct Submission

Submitted (05-SEP-2002) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

3 (bases 1 to 120926)

Buell, R.

Direct Submission

Submitted (28-SEP-2002) The Institute for Genomic Research, 9712

Medical Center Dr, Rockville, MD 20850, USA

On Sep 27, 2002 this sequence version replaced gi:22726029.

* NOTE: This is a 'working draft' sequence. It currently

* consists of 8 contigs. Gaps between the contigs

* are represented as runs of N. The order of the pieces

* is believed to be correct as given, however the sizes

* of the gaps between them are based on estimates that have

* provided by the submitter.

* This sequence will be replaced

* by the finished sequence as soon as it is available and

* the accession number will be preserved.

* 1 3442: contig of 3442 bp in length

* 3443 3542: gap of unknown length

* 3543 5560: contig of 2318 bp in length

* 5561 5960: gap of unknown length

* 5961 18145: contig of 12185 bp in length

* 18146 18245: gap of unknown length

* 18246 72353: contig of 54108 bp in length

* 72354 72453: gap of unknown length

* 72454 88544: contig of 16091 bp in length

* 88545 88644: gap of unknown length

* 88645 109570: contig of 20926 bp in length

* 109571 111770: contig of 2103 bp in length

* 111771 111870: gap of unknown length

* 111871 120326: contig of 9056 bp in length.

Location/Qualifiers

1..120926

/organism="Oryza sativa (japonica cultivar-group)"

/mol_type="genomic DNA"

/cultivar="Nipponbare"

/db_xref="taxon:39947"

/chromosome="11"

/clone="OSJNB0026K23"

/note="japonica cultivar-group"

BASE COUNT 34060 a 26488 c 25743 g 33915 t 720 others

ORIGIN

Query Match 3.9%; Score 26; DB 2; Length 120926;
Best Local Similarity 100.0%; Pred.No. 0.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 GCTCTGCTCTCTCTCTCTCTCTCTCTCT 557

Db 78382 GCTCTGCTCTCTCTCTCTCTCTCTCT 78407

RESULT 10
AC023549 155185 bp DNA linear HTG 10-SEP-2003
LOCUS Homo sapiens clone RP11-60C17, WORKING DRAFT SEQUENCE, 17 unordered
DEFINITION AC023549

pieces

AC023549

AC023549.3 GI:10047834

HTG; HTGS_PHASE1; HTGS_DRAFT.

SOURCE Homo sapiens (human)

ORGANISM Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 155185)

Birken, B., Linton, J., Nusbaum, C. and Lander, E.

HOMO SAPIENS CHROMOSOME, clone RP11-60C17

TITLE

JOURNAL
REFERENCE
AUTHORS

Unpublished
2 (bases 1 to 155185)
Barren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N., Anderson, S., Baldwin, J., Barna, N., Bida, F., Boguslavsky, L., Boukhalter, B., Brown, A., Burkett, G., Campopiano, A., Castle, A., Choepel, Y., Collangelo, M., Collins, S., Collymore, A., Cooke, P., DeArellano, K., Dewar, K., Dodge, S., Domino, M., Doyle, M., Fennestor, J., Ferreira, P., FitzHugh, M., Forrest, C., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Goyette, M., Graham, L., Grand-Pierre, N., Grant, G., Hagos, B., Heatford, A., Horton, L., Howland, J. C., Iliev, I., Johnson, R., Jones, C., Kann, L., Karatas, A., Klein, J., Landers, T., Largocque, K., Lehoczy, J., Levine, R., Lieu, C., Liu, G., Locke, K., Macdonald, P., Marquis, N., McCarthy, M., McEwan, P., McGurk, A., McKernan, K., McPheeters, R., Meldrum, J., Menes, L., Mihova, T., Miranda, C., Mlenga, V., Morrow, J., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neill, D., Oliver, F. W., Peterson, K., Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rogov, P., Rothman, D., Roy, A., Santos, R., Schauer, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J., Tirrell, A., Travers, M., Trigilio, J., Vassiliev, H., Viei, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zimmer, A. and Zody, M.

TITLE
JOURNAL

Submitted (15-FEB-2000) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
On Sep 10, 2000 this sequence version replaced gi:7331469.
All repeats were identified using RepeatMasker:

Sait, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>

----- Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIGR
Web site: <http://seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu

----- Project Information

Center project name: L6394

Center clone name: 60_C17

----- Summary Statistics

Sequencing vector: M13; X778.5; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 144294 bases at least Q40
Consensus quality: 148457 bases at least Q30
Consensus quality: 150798 bases at least Q20
Insert size: 148000; agarose-fp
Insert size: 153585; sum-of-ctiggs
Quality coverage: 5.0 in Q20 bases; agarose-fp
Quality coverage: 4.8 in Q20 bases; sum-of-ctiggs

NOTE: This is a 'working draft' sequence. It currently consists of 17 contigs. The true order of the pieces is not known and their order in this sequence record is arbitrary. Gaps between the contigs are represented as runs of N, but the exact sizes of the gaps are unknown. This record will be updated with the finished sequence, as soon as it is available and the accession number will be preserved.

1 1127: contig of 1127 bp in length
* 1128 1227: gap of 100 bp
* 1228 2232: contig of 1005 bp in length
* 2233 2332: gap of 100 bp
* 2333 3444: contig of 1112 bp in length
* 3445 3544: gap of 100 bp
* 3545 5242: contig of 1698 bp in length
* 5243 5342: gap of 100 bp
* 5343 8059: contig of 2717 bp in length
* 8060 8159: gap of 100 bp
* 8160 11253: contig of 3094 bp in length
* 11254 11353: gap of 100 bp
* 11354 14824: contig of 3471 bp in length
* 14825 14924: gap of 100 bp
* 14925 49854: contig of 34930 bp in length

* 49855 43954: gap of 100 bp
* 49856 53870: contig of 3316 bp in length
* 53871 53970: gap of 100 bp
* 53971 59531: contig of 5561 bp in length
* 59532 59631: gap of 100 bp
* 59632 65752: contig of 6121 bp in length
* 65753 65852: gap of 100 bp
* 65853 71898: contig of 6046 bp in length
* 71899 71998: gap of 100 bp
* 71999 79998: contig of 7998 bp in length
* 79999 80396: gap of 100 bp
* 80397 97717: contig of 17621 bp in length
* 97718 97817: gap of 100 bp
* 97818 120375: contig of 24258 bp in length
* 120376 122175: gap of 100 bp
* 122176 124445: contig of 30270 bp in length
* 124446 152545: gap of 100 bp
* 152546 155185: contig of 2640 bp in length.

FEATURES

Location/Qualifiers
source
1..155185
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/clone_lib="RPC1-11 Human Male BAC"

1..1127
/note="assembly_fragment"

misc_feature

1228..2232
/note="assembly_fragment"

misc_feature

2333..3444
/note="assembly_fragment"

misc_feature

3545..5242
/note="assembly_fragment"

misc_feature

5343..8059
/note="assembly_fragment"

misc_feature

8160..11253
/note="assembly_fragment"

misc_feature

11354..14824
/note="assembly_fragment"

misc_feature

14925..49854
/note="assembly_fragment"

misc_feature

49955..53870
/note="assembly_fragment"

misc_feature

53971..59531
/note="assembly_fragment"

misc_feature

59632..65752
/note="assembly_fragment"

misc_feature

65853..71898
/note="assembly_fragment"

misc_feature

71999..79998
/note="assembly_fragment"

misc_feature

80397..97717
/note="assembly_fragment"

misc_feature

97818..120375
/note="assembly_fragment"

misc_feature

122176..152445
/note="assembly_fragment"

misc_feature

152546..155185
/note="assembly_fragment"

misc_feature

clone_end:77
vector_side:right

BASE COUNT 47375 a 23006 c 28938 g 48262 t 1604 others
ORIGIN

Query Match 3.9%; Score 26; DB 2; Length 155185;
Best Local Similarity 100.0%; Pred. No. C.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 CTCGCGCCCTCCCTCTCCCTCTCC 248

DB 54935 CTCGCGCCCTCCCTCTCCCTCTCC 54960


```

RESULT 11
LOCUS       AC022931
DEFINITION Homo sapiens chromosome 5 clone RP11-356D23 map 5, WORKING DRAFT
SEQUENCE    17 unordered pieces.
AC022931
VERSION     AC022931.3 GI:7249435
KEYWORDS    HTG; HTGS_PHASE1; HTGS_DRAFT.
SOURCE      Homo sapiens
ORGANISM    Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1 (bases 1 to 157834)
AUTHORS     Birren, B., Linton, L., Nusbaum, C. and Lander, E.
TITLE       Homo sapiens chromosome 5, clone RP11-356D23
JOURNAL     Unpublished
REFERENCE   2 (bases 1 to 157834)
AUTHORS     Birren, B., Linton, L., Nusbaum, C., Lander, E., Abraham, H., Allen, N.,
            Anderson, S., Baldwin, J., Barna, K., Beckerly, R., Beda, F.,
            Boguski, M., Bouknight, B., Brown, A., Burkett, G., Castle, A.,
            Choquel, Y., Colangelo, M., Collins, S., Colymore, A., Cooke, P.,
            DeArillano, K., Dewar, K., Domino, M., Doyle, M., Fenster, J.,
            Ferreira, P., Fitzhugh, W., Forrest, C., Gage, D., Galagan, J.,
            Gardyna, S., Grant, G., Hagos, B., Heaford, A., Horton, L.,
            Howland, J.C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
            Howland, J.C., Johnson, R., Jones, C., Kann, L., Karas, A., Klein, J.,
            Landers, T., Lehotzky, J., Levine, R., Lien, C., Liu, G., Locke, K.,
            Macdonald, P., Marquis, N., McEwan, P., McGurk, A., McKernan, K.,
            McPheeters, R., Meidrum, J., Meneus, L., Morrow, J., Naylor, J.,
            Norman, C.H., O'Connor, T., O'Donnell, P., Olivari, T.M., Peterson, K.,
            Pierre, N., Pisani, C., Pollara, V., Raymond, C., Riley, R., Rotman, D.,
            Roy, A., Santos, R., Severy, P., Spencer, B., Stange-Thomann, N.,
            Stojanovic, N., Subramanian, A., Talamas, J., Testaye, S., Theodore, J.,
            Tirrell, A., Vassiliev, H., Vieri, R., Vo, A., Wu, X., Wyman, D., Ye, W.,
            Zimmer, A. and Zody, M.
            Direct Submission
            Submitted (07-FEB-2000) Whitehead Institute/MIT Center for Genome
            Research, 320 Charles Street, Cambridge, MA 02141, USA
            On Mar 16, 2000 this sequence version replaced GI:7008859.
            All repeats were identified using RepeatMasker:
            Smit, A.P.A. & Green, P. (1996-1997)
            http://ftp.genome.washington.edu/RW/RepeatMasker.html
            ----- Genome Center
            Center: Whitehead Institute/ MIT Center for Genome Research
            Center code: WIBR
            Web site: http://www-seq.wi.mit.edu
            Contact: sequence.submissions@genome.wi.mit.edu
            ----- Project Information
            Center project name: L6343
            Center clone name: 356D_23
            ----- Summary Statistics
            Sequencing vector: M13; M77815; 100% of reads
            Chemistry: Dye-terminator Big Dye; 100% of reads
            Assembly program: Phrap; version 0.960731
            Consensus quality: 148896 bases at least Q40
            Consensus quality: 153119 bases at least Q30
            Consensus quality: 154676 bases at least Q20
            Insert size: 160000; agarose-fp
            Insert size: 156234; sum-of-contigs
            Quality coverage: 4.8 in Q20 bases; agarose-fp
            Quality coverage: 4.9 in Q20 bases; sum-of-contigs
            -----
            * NOTE: This is a 'working draft' sequence. It currently
            * consists of 17 contigs. The true order of the pieces
            * is not known; and their order in this sequence record is
            * arbitrary. Gaps between the contigs are represented as
            * runs of N, but the exact sizes of the gaps are unknown.
            * This record will be updated with the finished sequence.
            * as soon as it is available and the accession number will
            * be preserved.
            *
            * 1 1404: contig of 1404 bp in length
            * 1405 1504: gap of 100 bp
            * 1505 4290: contig of 2766 bp in length
            * 4291 4390: gap of 100 bp
            *
            * 4391 7101: contig of 2711 bp in length
            * 7102 7421: gap of 100 bp
            * 7422 10425: contig of 3224 bp in length
            * 10426 10525: gap of 100 bp
            * 10526 14004: contig of 3479 bp in length
            * 14005 14104: gap of 100 bp
            * 14105 18152: contig of 4048 bp in length
            * 18153 22230: gap of 100 bp
            * 22231 22310: contig of 3978 bp in length
            * 22311 28331: contig of 5901 bp in length
            * 28332 34381: contig of 6053 bp in length
            * 34382 34481: gap of 100 bp
            * 34482 40840: contig of 6359 bp in length
            * 40841 40940: gap of 100 bp
            * 40941 48630: contig of 7690 bp in length
            * 48631 48730: gap of 100 bp
            * 48731 60384: contig of 11654 bp in length
            * 60385 60484: gap of 100 bp
            * 60485 72266: contig of 11782 bp in length
            * 72267 95070: contig of 12704 bp in length
            * 95071 85170: gap of 100 bp
            * 85171 130748: contig of 15578 bp in length
            * 130749 130848: gap of 100 bp
            * 130849 121955: contig of 21107 bp in length
            * 121956 122055: gap of 100 bp
            * 122056 157834: contig of 35779 bp in length.
            *
            * Location/Qualifiers
            * source
            * 1..157834
            * /organism="Homo sapiens"
            * /mol_type="genomic DNA"
            * /db_xref="taxon:9606"
            * /chromosome="5"
            * /map="5"
            * /clone="RP11-356D23"
            * /clone_lib="RPCr-11 Human Male BAC"
            * 1..1404
            * /note="assembly_fragment"
            * clone_end:77
            * vector_side:right
            * misc_feature
            * 1505..4290
            * /note="assembly_fragment"
            * misc_feature
            * 4391..7101
            * /note="assembly_fragment"
            * misc_feature
            * 7202..10425
            * /note="assembly_fragment"
            * misc_feature
            * 10526..14004
            * /note="assembly_fragment"
            * misc_feature
            * 14105..18152
            * /note="assembly_fragment"
            * misc_feature
            * 18253..22230
            * /note="assembly_fragment"
            * misc_feature
            * 22311..28331
            * /note="assembly_fragment"
            * misc_feature
            * 28332..34381
            * /note="assembly_fragment"
            * misc_feature
            * 34482..40840
            * /note="assembly_fragment"
            * misc_feature
            * 40941..48630
            * /note="assembly_fragment"
            * misc_feature
            * 48731..50384
            * /note="assembly_fragment"
            * misc_feature
            * 60485..72266
            * /note="assembly_fragment"
            * misc_feature
            * 72267..85070
            * /note="assembly_fragment"
            * misc_feature
            * 85171..130748
            * /note="assembly_fragment"
            * clone_end:596
            * vector_side:left"
            * 130849..1521955
            *
            * misc_feature
            * 130849..1521955

```

```

misc_feature      /note="assembly_fragment"
122056..157834
/note="assembly_fragment"
BASE COUNT      46053 a 31958 c 29611 g 48507 t 1605 others
ORIGIN

Query Match      3.9%; Score 26; DB 2; Length 157834;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 CTCGCGCCCTCTCCCTCTCCCTCTCC 248
|||||
Db 6178 CTCGCGCCCTCTCCCTCTCCCTCTCC 6203
|||||

RESULT 12
AC114969      164526 bp      DNA      linear      PRI 01-JUN-2002
DEFINITION      Homo sapiens chromosome 5 clone RP11-356D23, complete sequence.
ACCESSION      AC114969
VERSION      AC114969.2 GI:21306652
KEYWORDS      HTG.
SOURCE      Homo sapiens (human)
ORGANISM      Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE      1 'bases 1 to 164526'
AUTHORS      DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL      Direct Submission
UNPUBLISHED
REFERENCE      2 (bases 1 to 164526)
AUTHORS      DOE Joint Genome Institute.
JOURNAL      Direct Submission
SUBMITTED (14-MAR-2002) Production Sequencing Facility, DOE Joint
Genome Institute, 2800 Mitchell Drive, Walnut Creek, CA 94598, USA
REFERENCE      3 (bases 1 to 164526)
AUTHORS      DOE Joint Genome Institute and Stanford Human Genome Center.
JOURNAL      Direct Submission
SUBMITTED (01-JUN-2002) DOE Joint Genome Institute, 2800 Mitchell
Drive, Walnut Creek, CA 94598, USA
COMMENT      On Jun 1, 2002 this sequence version replaced gi:19424430.
Draft Sequence Produced by DOE Joint Genome Institute
www.jgi.doe.gov
Finishing Completed at Stanford Human Genome Center
www.sngc.stanford.edu
Quality: Phrap Quality >=40 99.8% of Sequence;
Estimated Total Number of Errors is 0.2.
FEATURES
source
1..164526
/organism="Homo sapiens"
/mol_type="genomic DNA"
/db_xref="taxon:9606"
/chromosome="5"
/clone="RP11-356D23"
BASE COUNT      52577 a 32245 c 32724 g 46980 t
ORIGIN

Query Match      3.9%; Score 26; DB 9; Length 164526;
Best Local Similarity 100.0%; Pred. No. 0.026;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 CTCGCGCCCTCTCCCTCTCCCTCTCC 248
|||||
Db 94483 CTCGCGCCCTCTCCCTCTCCCTCTCC 94508
|||||

RESULT 13
BX293990/c
LOCUS      BX293990
DEFINITION      Mus musculus chromosome 2 clone RP24-222N16, *** SEQUENCING IN
PROGRESS ***.
ACCESSION      BX293990
VERSION      BX293990.5 GI:30348793

KEYWORDS      HTG; HTGS PHASE2; HTGS ACTIVEFIN; HTGS_DRAFT; HTGS_FULLTOP.
Mus musculus (house mouse).
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 183011)
Johnson, C.
Direct Submission
Submitted (01-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB0 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk
On May 2, 2003 this sequence version replaced gi:30230876.
Sequence from the Mouse Genome Sequencing Consortium whole genome
shotgun may have been used to confirm this sequence. Sequence data
from the whole genome shotgun alone has only been used where it has
a phred quality of at least 30.
----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk
----- Project Information
Center project name: EN22N16
----- Summary Statistics
Assembly program: XGAP4; version 4.5
Chemistry: Dye-terminator; 10% of reads
Consensus quality: 18293 bases at least Q40
Consensus quality: 18253 bases at least Q36
Consensus quality: 18265 bases at least Q20
Insert size: 18301; sum-of-contigs
Quality coverage: 8.90x in Q20 bases; sum-of-contigs Quality
coverage: 9.70x in Q20 bases; agarose-fp
-----
* NOTE: This is a 'working draft' sequence. It currently
* consists of 1 contigs. Gaps between the contigs
* are represented as runs of N. The order of the pieces
* is believed to be correct as given, however the sizes
* of the gaps between them are based on estimates that have
* provided by the submitter.
* This sequence will be replaced
* by the finished sequence as soon as it is available and
* the accession number will be preserved.
* 183011: contig of 183011 bp in length.
* Location/Qualifiers
1..183011
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP24-222N16"
/clone_id="RPC1-24"
1..183011
/note="assembly_fragment:02752"
clone_end:17
vector_side:left"
BASE COUNT      50069 a 41722 c 40049 g 51171 t
ORIGIN

Query Match      3.9%; Score 26; DB 2; Length 183011;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 223 CTCGCGCCCTCTCCCTCTCCCTCTCC 248
|||||
Db 162078 CTCGCGCCCTCTCCCTCTCCCTCTCC 162053
|||||

RESULT 14
AC129216/c
LOCUS      AC129216
DEFINITION      Mus musculus chromosome 14 clone RP24-446C11, complete sequence.
ACCESSION      AC129216

```

```

VERSION AC129216.4 GI:30061471
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 190338)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE The sequence of Mus musculus clone
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 190338)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (27-JUL-2002) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 3 (bases 1 to 190338)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (20-MAR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
REFERENCE 4 (bases 1 to 190338)
AUTHORS McPherson,J.D. and Waterston,R.H.
TITLE Direct Submission
JOURNAL Submitted (23-APR-2003) Genome Sequencing Center, 4444 Forest Park
Parkway, St. Louis, MO 63108, USA
COMMENT On Apr 23, 2003 this sequence version replaced gi:29126529.

----- Genome Center -----
Center: Washington University Genome Sequencing Center
Center code: WUGSC
Web site: http://genome.wustl.edu/gsc/index.shtml
Contact: submissions@wustl.wustl.edu
----- Project Information -----
Center project name: M_BB0446C11
----- Location/Qualifiers -----
1..190338
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="14"
/clone="RP24-446C11"
BASE COUNT 52061 a 39827 c 41432 g 58018 t
ORIGIN
Query Match 3.9%; Score 26; DB 10; Length 190338;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 223 CTCGCGCCCTCTCCCTCTCCCTCTCC 248
DB 141919 CTCGCGCCCTCTCCCTCTCCCTCTCC 141894

RESULT 15
AL928914/c 199873 bp DNA linear ROD Cl-JUN-2003
LOCUS Mouse DNA sequence from clone RP23-399J8 on chromosome 2, complete
sequence.
DEFINITION AL928914
ACCESSION AL928914.5 GI:31335573
VERSION AL928914
KEYWORDS HTG.
SOURCE Mus musculus (house mouse)
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 199873)
AUTHORS Humphries,M.
TITLE Direct Submission
JOURNAL Submitted (30-MAY-2003) Wellcome Trust Sanger Institute, Hinxton,
Cambridgeshire, CB10 1SA, UK. E-mail enquiries:
humquery@sanger.ac.uk Clone requests: clonerequests@sanger.ac.uk
COMMENT On Jun 2, 2003 this sequence version replaced gi:29886622.

```

Sequence from the Mouse Genome Sequencing Consortium whole genome shotgun may have been used to confirm this sequence. Sequence data from the whole genome shotgun alone has only been used where it has a phred quality of at least 30.

----- Genome Center
Center: Wellcome Trust Sanger Institute
Center code: SC
Web site: http://www.sanger.ac.uk
Contact: humquery@sanger.ac.uk

During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.

This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one Y13 subclone; and the assembly was confirmed by restriction digest, except on the rare occasion of the clone being a YAC.

The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: EM, EMBL; SW, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at <http://www.sanger.ac.uk/projects/Celegans/wormpep> RP23-399J8 is from the RP23-399J8 BAC library constructed by the group of Pieter de Jong. For further details see <http://www.chori.org/bacpac/home.htm>

FEATURES
Location/Qualifiers
Source
1..199873
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
/chromosome="2"
/clone="RP23-399J8"
/clone_lib="RPCI-23"

BASE COUNT 50089 a 47425 c 46975 g 55384 t
ORIGIN

Query Match 3.9%; Score 26; DB 10; Length 199873;
Best Local Similarity 100.0%; Pred. No. 0.025;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

CY 223 CTCGCGCCCTCTCCCTCTCCCTCTCC 248
DB 89873 CTCGCGCCCTCTCCCTCTCCCTCTCC 89848

Search completed: November 17, 2003, 21:16:26
Job time : 2571 secs